REMARKS

The present application was filed on June 30, 1999 with claims 1 through 35. Claims 1 through 35 are presently pending in the above-identified patent application.

In the Office Action, the Examiner objected to the cited reference IDS 4 because of lacking English description/translation. The Examiner also rejected claims 1-5, 8, 10-14, 16-19, 21-26 and 28-35 under 35 U.S.C. §102(b) as being anticipated by Chen et al., "Speaker, Environment and Channel Change Detection and Cluster via the Bayesian Information Criterion," Proc. of the DARPA Broadcast News Workshop (Feb. 1998), hereinafter, referred to as "Chen." In addition, claims 6-7, 9, 20 and 27 were rejected under 35 U.S.C. §103(a) as being unpatentable over Chen in view of well known prior art and claim 15 was rejected under 35 U.S.C. §103(a) as being unpatentable over Chen in view of Kleider et al. (United States Patent No. 5,930,748).

The present invention automatically identifies speakers in an audio source by concurrently segmenting the audio source and clustering the segments corresponding to the same speaker.

Information Disclosure Statement

The Examiner notes that the references listed in the IDS submitted on April 25, 2007 have been considered except for the "cited reference IDS 4 that cannot be considered because of lacking [SIC] English description/translation."

Applicants note that 37 CFR 1.98(a) states that any information disclosure statement filed under Sec. 1.97 shall include the items listed in paragraphs (a)(1), (a)(2) and (a)(3) of this section. 37 CFR 1.98(a)(3)(i) allows for a *concise explanation of the relevance, as it is presently understood by the individual designated in Sec. 1.56(c) most knowledgeable about the content of the information, of each patent, publication, or other information listed that is not in the English language.* The concise explanation may be either separate from applicant's specification or incorporated therein. According to 37 CFR 1.98(a)(3)(ii), a copy of the translation is *only required if* a written English-language translation of a non-English-language document, or portion thereof, is within the possession, custody, or control of, or is readily available to any individual designated in Sec. 1.56(c). Applicants also note that 37 CFR 1.56(c) *designates each attorney or agent who prepares or prosecutes the application,* and note that a written English-language

translation of the Japanese document, or portion thereof, was not within the possession, custody, or control of any individual designated in Sec. 1.56(c), nor was it readily available.

Thus, the Statement of Relevance and the related Japanese patent abstract publication included in the IDS dated April 25, 2007 should be considered by the Examiner. Applicants respectfully request that the Examiner consider the cited document.

Independent Claims 1, 16, 23, and 30-35

Independent claims 1, 16, 23 and 30-35 are rejected under 35 U.S.C. §102(b) as being anticipated by Chen et al.

In the Office Action dated August 27, 2002, the Examiner asserted that Chen discloses speaker, environment and channel change detection and clustering via the Bayesian Information Criterion for segmenting the audio stream into homogeneous regions according to speaker identity, environmental condition and channel condition and clustering speech segments into homogeneous clusters according to speaker identity, environmental condition and channel (citing page 1, paragraph 2) which reads on the claimed "method of tracking a speaker in an audio source, said method comprising the steps of identifying potential segment boundaries in said audio source; and clustering homogeneous segments from said audio source substantially concurrently with said identifying step."

In the Response to Office Action dated December 26, 2002, Applicants submitted that, while Chen discloses segmenting an audio stream into homogeneous regions and clustering speech segments into homogeneous clusters, the audio stream is *first* segmented and *then* clustered. Applicants noted, as further evidence that the clustering in Chen is performed only after the audio stream has been segmented, that Section 4.1 indicates that each segment is compared to all other segments before clustering is finalized. In addition, Section 4.2, first paragraph indicates that the data set consists of an audio file that has been "hand-segmented into 824 short segments."

In the Office Action dated March 7, 2003, the Examiner notes that the prior art cites that "our segmentation algorithm can successfully detect acoustic changes" (Chen: abstract) and that "we first examine whether our detected change points were

true." (Chen: Section 3.3, paragraph 3.) The Examiner asserts that this suggests that Chen not only employs its own segmenting mechanism, but is also capable of combining segmentation with clustering "substantially concurrently."

The Examiner also asserts that Chen suggests that clustering does not need completely segmented data, such that a clustering process may be combined with a segmenting process together substantially concurrently, since Chen discloses that "it is also clear that our criterion can be applied to top-down methods." (Chen: Section 4.1, paragraph 4.)

The Examiner further asserts that a clustering step can be inserted in the segmentation loop, in Chen, Section 3.2, paragraph 1, and that Chen is capable of combining segmentation and clustering since the segmentation and clustering algorithms are based on the BIC algorithm and since equations (2), (3), and (8) have no limitation for combining segmentation and clustering.

Applicants acknowledge that Chen employs its own segmenting mechanism, but find no indication of or suggestion to perform segmentation and clustering "substantially concurrently" in the cited text. Applicants note that the Examiner asserts that Chen is *capable of* this, but does not assert that Chen suggests or discloses combining segmentation with clustering substantially concurrently.

Applicants also note that, in the top-down method, a hypothesis is made regarding the number of clusters. Then, a test is made to determine if the number of clusters hypothesized actually "fits" the data. Alternatively, in the bottom-up method, the number of clusters is determined from the data. Thus, the capability to utilize a top-down method does not suggest that segmentation is performed substantially concurrently with the clustering process.

Regarding the final assertion made by the Examiner, Applicants also note that, whether or not Chen is *capable* of combining segmentation and clustering, there is no disclosure or suggestion to do so.

Thus, Chen does not disclose or suggest a "method of tracking a speaker in an audio source, said method comprising the steps of identifying potential segment boundaries in said audio source; and clustering homogeneous segments from said audio source substantially concurrently with said identifying step," as required by independent

claims 1, 16, 30, 31, 32 and 33 of the present invention. Similarly, independent claims 23, 34 and 35 require that the segmentation and clustering are performed on the "same pass" through said audio source.

Examiner's Answer

In the Examiner's Answer dated December 17, 2003, the Examiner states that it is believed that the limitation "substantially concurrently" has no patentable weight, because the Applicant does not have any clear definition and/or description in the claim or in the specification about this limitation, and does not give any conditions to apply this limitation. The Examiner also asserts that the prior art explicitly and/or implicitly discloses all the limitations regarding claim 1, including the limitation of "substantially concurrently," based on the interpretation of the claim language and the understanding (of the) prior art teachings. In particular, the Examiner asserts that the performance of the two steps (segmentation and clustering) may be associated with many time related factors, including computing speed, simple rate, and total stream size.

The Examiner further asserts that the fact that the clustering in Chen is performed only after the audio stream has been segmented and that each segment is compared to all other segments before clustering is finalized is not relevant to claim 1 since claim 1 does not recite these limitations.

The Examiner also notes that one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references.

Regarding the Examiner's assertion that the limitation "substantially concurrently" has no patentable weight, Applicants note that the word "substantially" has a well known and well understood definition in claim language. Its meaning is sufficiently clear in the teachings of the specification such that a person of ordinary skill in the art would understand the limitation without the need to apply conditions.

Regarding the Examiner's assertion that the prior art explicitly and/or implicitly discloses all the limitations regarding claim 1, including the limitation of "substantially concurrently," based on the interpretation of the claim language and the understanding (of the) prior art teachings, Applicants note that the broad interpretations made by the Examiner are ***not consistent*** with the specification and are ***not consistent***

with the interpretation of the specification that a person of ordinary skill in the art would make. For example, the interpretation that "substantially concurrent" is related to a computing speed is not disclosed or suggested by the specification and the cited term would not be interpreted in this manner by a person of ordinary skill in the art. As disclosed on page 2 (lines 16-26) of the original specification, "the present invention concurrently *segments an audio file and clusters the segments* corresponding to the same speaker." Thus, the term "substantially concurrent" is related to the *parallel execution* of the segmentation and clustering steps. See, also, FIG. 2.

More specifically, Applicants note that these limitations are clearly captured in claim 1, which recites the limitations of identifying potential segment boundaries in said audio source; and clustering homogeneous segments from said audio source substantially concurrently with said identifying step. Claim 1 requires the clustering of homogeneous segments *substantially concurrently* with said identifying step. Chen, therefore, actually teaches away from the present invention by teaching that the clustering is performed only after the audio stream has been segmented. Thus, contrary to the Examiner's assertion, the limitations cited by the Examiner in reference to Chen are ***clearly relevant*** to the consideration of claim 1.

Applicants also note that the references were not attacked individually, but were reviewed to demonstrate that ***none*** of the references contain a limitation required by the claims of the present invention and that, therefore, the prior art does not pose a bar to patentability.

Additional Cited References

Kleider et al. was also cited by the Examiner in rejecting claim 15 for its disclosure that the information of the speaker model data may include a speaker name. Applicants note that the inventors listed in United States Patent Number 5,157,763 (referred to by the Examiner in the Final Office Action) are not Kleider et al. Applicants did find, however, United States Patent Number 5,930,748 in the Notice of References Cited and respond to that reference below.

Applicants note that Kleider et al. is directed to a "method of identifying an individual from a predetermined set of individuals using a speech sample spoken by the individual. The speech sample comprises a plurality of spoken utterance, and each

individual of the set has predetermined speaker model data." Cited, Summary of the Invention. Kleider et al. do not address the issue of segmenting speech.

Thus, Kleider et al. do not disclose or suggest a "method of tracking a speaker in an audio source, said method comprising the steps of identifying potential segment boundaries in said audio source; and clustering homogeneous segments from said audio source substantially concurrently with said identifying step," as required by independent claims 1, 16, 30, 31, 32 and 33 of the present invention. Similarly, independent claims 23, 34 and 35 require that the segmentation and clustering are performed on the "same pass" through said audio source.

Dependent Claims 2-15, 17-22 and 24-29

Dependent claims 2-5, 8, 10-14, 17-19, 21-22, 24-26, and 28-29 were rejected under 35 U.S.C. §102(b) as being anticipated by Chen et al., claims 6-7, 9, 20 and 27 were rejected under 35 U.S.C. §103(a) as being unpatentable over Chen in view of well known prior art, and claim 15 was rejected under 35 U.S.C. §103(a) as being unpatentable over Chen in view of Kleider et al.

Claims 2 through 15, 17 through 22, and 24 through 29 are dependent on claims 1, 16 and 23, respectively, and are therefore patentably distinguished over Chen and Kleider et al., alone or in combination with well known prior art, because of their dependency from independent claims 1, 16 or 23 for the reasons set forth above, as well as other elements these claims adds in combination to their base claim.

All of the pending claims, i.e., claims 1-35, are in condition for allowance and such favorable action is earnestly solicited.

If any outstanding issues remain, or if the Examiner has any further suggestions for expediting allowance of this application, the Examiner is invited to contact the undersigned at the telephone number indicated below.

The Examiner's attention to this matter is appreciated.

Respectfully submitted,

Date: October 11, 2007

Kevin M. Mason
Attorney for Applicants
Reg. No. 36,597
Ryan, Mason & Lewis, LLP
1300 Post Road, Suite 205
Fairfield, CT 06824
(203) 255-6560